Learning academic formulaic sequences

To appear in ***Journal of English for Academic Purposes***

Elke Peters

KU Leuven

Faculty of Arts

Campus Antwerp

Sint-Andriesstraat 2

B-2000 ANTWERP

Elke.peters@arts.kuleuven.be

Paul Pauwels

KU Leuven

Faculty of Arts

Campus Antwerp

Sint-Andriesstraat 2

B-2000 ANTWERP

Paul.pauwels@arts.kuleuven.be

Abstract

This paper reports on a classroom-based study that explored the effect of explicit, vocabulary-focused instruction on English as a Foreign Language (EFL) students' recognition, cued output and spontaneous use of academic formulaic sequences (FS). In addition, the study aimed to shed some light on which type of classroom activity might be most beneficial. Data were collected among second-year EFL business students (L1 = Dutch) in a classroom-based experiment during students' regular English classes. A pre-test post-test within-subject design was adopted. Twenty-four pedagogically relevant FS were selected and offered in three types of activities: 1) recognition activities, 2) cued output activities, or 3) a combination of recognition and cued output activities. Learning gains were measured in a recognition test, a cued output test, a writing test and students' end-of-year assignment (= spontaneous productive use). The findings revealed that students made significant learning gains from pretest to posttest. In addition, analyses of students' end-of-year assignments showed that students spontaneously used a considerable number of FS in their assignment. Finally, the results tentatively suggest that activities including cued output resulted in higher learning gains than recognition activities.

Introduction

Corpus studies have revealed the omnipresence and importance of formulaic sequences (FS) in academic writing. FS are frequent "combinations of at least two words favored by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalization" (Erman & Warren, 2000, p. 31). In academic writing, they are used to realize functional units of academic genres. Their use allows writers to signal stages in their discourse, to display the appropriate level of formality and to meet the expectations of the academic community (Coxhead & Byrd, 2007; Jones & Haywood, 2004). Mastery of academic FS is crucial if one wants to succeed as an academic writer. However, English as a Foreign Language (EFL) learners struggle with the appropriate use of such sequences, making their academic prose come across as inappropriate and too chatty (Gilquin & Paquot, 2008). Learners do not seem to acquire FS incidentally, even if they know individual items comprising them, which may suggest that an explicit learning approach is called for (Meunier, 2012). However, few studies on practical applications have been published so far (Hyland, 2012).

In order to address this gap, the current study explores the effect of vocabulary-focused instruction on EFL learners' recognition, cued output and spontaneous use of academic FS. An ecologically valid, classroom-based experiment, resembling a real-life learning experience, was set up, in which learners engaged in explicit vocabulary activities focusing on academic FS. Learners' progress was measured from pretests to posttests. The aim of this study is to shed some light on whether it is beneficial to devote classroom time to the explicit teaching of a limited number of academic FS and how this might be best achieved.

Formulaic sequences

Corpus research has shown that language is to a large extent formulaic in nature (Meunier, 2012; Sinclair, 1991). Formulaic sequences (FS) are important not only because they are ubiquitous in language but also because they are necessary for appropriate, fluent language production and comprehension (Schmitt & Carter, 2004; Meunier, 2012). As a result, their use can be considered a marker of proficiency of foreign language learners. It has been advocated that FL learners need to acquire a considerable number of FS in order to become fluent and proficient FL speakers. Cowie (1992, p. 10) even argues that "it is impossible to perform at a level acceptable to native users, in writing or in speech, without controlling an appropriate range of multiword units". Although a lack of FS may still result in meaningful, grammatically correct language output, it will make FL learners sound odd and non-native-like.

The same holds true for spoken and written academic discourse (Coxhead, 2008; Hyland, 2008, 2012). Academic FS, such as *on the other hand* or *a central issue*, are used to signal stages and to realize functional units of this register (Cortes, 2013; Li & Schmitt, 2009). In this context, Hüttner (2007, p.97) distinguishes "genre-specific FS" and "genre-functional FS" (= FS used to realize a particular genre move) and argues that the latter are especially important for L2 learners because their appropriate use marks a writer as an insider in the academic discourse community. Let us illustrate this with a concrete example. Cortes (2013)[1] found that *studies have shown that* is almost exclusively found when previous literature is reviewed. In addition, sequences such as *the aim of this study is* or *the purpose of the present study was to* were only found in the move "to announce present research". Such sequences can thus be considered "defining markers of fluent writing and are important for the development of writing that fits the expectations of readers in academia" (Coxhead & Byrd, 2007, p. 135). Likewise, a lack of such sequences in students' writing may mark a writing assignment as inappropriate and too colloquial (Coxhead & Byrd, 2007; Gilquin &

Paquot, 2008; Hyland, 2012; Jones & Haywood, 2004). The development of lists containing the most frequently used academic FS (also called lexical bundles, lexical formulas, or multi-word constructions) is therefore not surprising (Ackermann & Chen, 2013; Biber, 2009; Cortes, 2013; Hyland, 2008; Liu, 2012; Simpson-Vlach & Ellis, 2010).

The importance and omnipresence of FS in academic writing means that mastering academic FS becomes a prerequisite for any FL learner who wants to be successful in their academic writing. FL learners should not only know how a text is organized in terms of functional units but also how these units are realized linguistically and lexically (Cortes, 2013). However, FL learners' use of FS tends to differ from that of native speakers (Ädel & Erman, 2012; Chen & Baker, 2010; Durrant & Mathews-Aydınlı, 2011). Chen and Baker (2010) demonstrated that the use of FS is linked to writing proficiency and that compared to native speakers, FL learners tend to favor certain FS but use fewer hedging devices. Ädel and Erman (2012) found that native speakers do not only use more FS, they also use a wider range of FS compared to FL learners. In addition, some FS were only found in native-speaker writing, whereas others were only identified in FL learners' writing. In general, non-native speaker writing contained more informal FS (e.g. *to find out*). Gilquin and Paquot (2008) also revealed learners' overuse of spoken-like features in academic prose.

Given the formulaic nature of academic writing and the challenges FL learners face in using FS appropriately, students should be familiarized with the FS typically associated with the communicative functions of academic writing (Hüttner, 2007). Jones and Haywood (2004) list the following advantages of using FS. They help students 1) meet the expectations of the academic community, 2) signal stages in their discourse, 3) express their ideas more economically, and 4) display the necessary level of formality. Coxhead and Byrd (2007) add that the use of FS makes the task of writing academic English easier because the writer is working with ready-made phrases rather than having to create each sentence word by word.

Teaching formulaic sequences

Boers and Lindstromberg (2012) provide an exhaustive review of intervention studies that have investigated which enhancement techniques have the potential to boost students' knowledge of FS. Although awareness-raising activities, such as Lewis' (1993) lexical approach, seem to promote students' use of FS, they may not always foster deep processing of the FS (Boers, Eyckmans, Kappel, Stengers & Demecheleer, 2006). However, the following factors do seem to have a beneficial effect: visual salience of the FS in the input (Peters, 2012), frequency of occurrence (Peters, 2014; Webb, Newton, & Chang, 2013), and access to meaning via glosses or dictionaries (Laufer, 2011; Peters, 2009, 2012). Boers and Lindstromberg have also consistently found positive, long-term effects of having students exploit formal properties of FS, e.g. by asking students to pay attention to sound repetition (alliteration, assonance) (Boers & Lindstromberg, 2005; Lindstromberg & Eyckmans, 2014). Other factors that have been found to contribute to students' learning FS are translation activities (Laufer & Girsai, 2009) and explicit vocabulary activities (Peters, 2014; Webb & Kagimoto, 2009). However, none of these studies centered around academic FS. The remainder of this section will be devoted to two studies which focused explicitly on the acquisition of academic FS.

Jones and Haywood (2004) carried out an exploratory study in the course of a ten-week EAP class to investigate whether explicit vocabulary instruction would affect students' awareness, cued production and free production of academic FS. Twenty students were involved in their study. The treatment group (N = 10) engaged in reading activities aiming at increasing students' awareness of FS (reading texts, concordances, finding equivalents, …), and in writing activities (four essays), whereas the control group (N = 11) did not. A pretest-posttest design was adopted to determine whether there were any learning gains. Jones and

Haywood found that the majority of students in the treatment group could recognize more FS in the reading posttest than in the pretest. Still, only a few students made progress in terms of production as measured by a C-test. Overall, Jones and Haywood did not find any improvement in students' essays. However, as the authors themselves pointed out, their study suffered from a number of methodological flaws. First, Jones and Haywood did not use the same pretests and posttests. Given the difference in items that were tested in the pretest and posttest, the results are not directly comparable. Furthermore, only seven participants took the C-test before and after the treatment, making it difficult to generalize any findings. Moreover, there were only two weeks between the two writing tests, which may not have been enough of a gap to discover any learning gains. Finally, it is not clear how often each single FS was offered in which activities. As a result, the type of activity as well as the frequency of occurrence may have been a confounding variable.

Li and Schmitt (2009) conducted a longitudinal case study in which they analyzed the FS of a Chinese-speaking student during her one-year MA program. Her writing assignments were analyzed in terms of the number and the appropriateness of the FS used. In addition, she was interviewed after each assignment about the sources of learning and the confidence with which she had used the FS in her writing. The findings revealed that she was able to learn a considerable number of new FS and to enhance mastery of sequences she already (partially) knew. She also gained confidence in using the FS. Regarding learning strategies, this student indicated that both implicit (academic reading) and explicit approaches had helped her in learning (new) FS. However, Li and Schmitt did not find a relationship between the number and appropriacy of FS used and her essay marks, which they partly attribute to the multiple assessment criteria that had been used making it difficult to single out one factor such as FS.

However, in spite of the increasing number of studies on foreign language learners' use and acquisition of FS in general, there is still an urgent need for more empirical and

classroom-based studies that explore whether and how students' knowledge and use of academic FS can be enhanced through vocabulary-focused instruction.

Aim and research questions

Given the importance of FS for academic writing and the scarcity of studies investigating the acquisition of academic FS, this study sets out to investigate whether vocabulary-focused instruction, in which the FS are the explicit focus of teaching, has the potential to boost students' awareness and use of academic FS. Research into the effectiveness of vocabulary teaching methods is important because it allows us to determine whether it is indeed worthwhile to devote classroom time to explicitly teaching (a number of) academic FS and how this should be realized.

The first purpose of this study is to investigate whether a vocabulary-focused approach adopted in a real classroom setting has the potential to boost foreign language learners' awareness and use of academic formulaic sequences. The second purpose is to shed some light on how this might best be achieved.

The following two sets of research questions were addressed:

1. Does vocabulary-focused instruction have an effect on (1) the number of FS recognized, (2) the number of FS recalled, and (3) the number of FS used spontaneously?
2. Does the type of form-focused instruction have an effect on (1) the number of FS recognized, (2) the number of FS recalled, and (3) the number of FS used spontaneously?

Method

The present study draws on Jones and Haywood's (2004) study since it investigates the effect of vocabulary-focused instruction of academic FS on students' recognition of FS, cued output

of FS and spontaneous use of FS in their writing. As in Jones and Haywood's study, an ecologically valid (= resembling a real-life, authentic learning situation), classroom-based study was set up, in which a pretest-posttest design was adopted to determine any learning gains. Data were collected in intact classes in the course of the academic year. Although the present study can thus be considered a replication study, it also differs from Jones and Haywood's study in a number of respects. First, the present study also aims to investigate which type of instruction is the most beneficial. Unlike the study by Jones and Haywood, the present study adopted a within-subject design that was implemented over a five-week period. In addition, our study explicitly aims to take into account some of the methodological issues that have been raised regarding their study:

- The pretests and posttests are identical to ensure that results are directly comparable;
- All target items are tested in the pretests and posttests;
- All pretests are administered prior to the treatment; all posttests after completion of the treatment;
- Frequency of occurrence of the target items is held constant;

Participants

The participants in this study were 29 Dutch-speaking EFL learners recruited from two parallel second-year business English classes at a Flemish, Dutch-medium university[2] in Belgium. All second-year Business Studies students have a compulsory Business English course in their program. Participants had studied English for at least five years prior to the treatment. Their proficiency level can be considered B1 to B2 of the Common European Framework of Reference. Only data from those participants who were present during the pre-test session, the post-test session and at least two of the three learning sessions were taken into account. Students were considered to be familiar with the concept of FS since the

business English course adopts Lewis's (1993) lexical approach and devotes a considerable amount of time to business-English collocations and FS.

Target items

Twenty-four FS were selected from the *Academic Phrasebank from the University of Manchester* (http://www.phrasebank.manchester.ac.uk) as target items (see Table 1). The following criteria were taken into account in the selection procedure:

- Each target item consists of at least two open-class constituent words[3]
- Each target item contains at least one word of Coxhead's (2000) Academic Word List (AWL) (i.e. the word in bold in Table 1)
- Each target item is listed in *The Corpus of Contemporary American English* (COCA) (Davies, 2008-)
- Each target item is pedagogically relevant, i.e. relevant to these students' end-of-year-writing assignment (see below), because it was a classroom-based experiment.

The website *The Academic Phrasebank from the University of Manchester* lists FS that might be useful for academic writing. All FS, which are generic in nature, are organized according to the main sections of a research paper. The website was specifically developed for non-native speakers of English. It was deemed appropriate for the selection of the target items because it allowed us to provide students with an easy-to-use and accessible source of academic FS, which could help them find and discover more FS than would be focused upon in the treatment. Given the nature and organization of students' end-of-year writing assignment (= a collaborative research paper), this website appeared to be particularly useful.

Table 1

List of target items

Recognition items	**Cued Output items**	**Recognition + Cued Output items**
a central **issue**	little **research** into	a **considerable** amount of (literature)
findings are **consistent** with	recent **evidence** suggests	**evidence** ... is **inconclusive**
findings **reveal**	findings/results **indicate**	studies **demonstrate**
the main **objective** of	... results **emerge** from ...	**Seek(s)** to investigate
to adopt an **approach**	this paper/study **focuses** on...	to collect **data**
to undertake **research**	to address **research** questions	questionnaire **consists**/ed of
to **conduct** an experiment	to administer a **survey**	to **conduct** an interview
to draw **conclusions**	to **obtain data**	to **select participants**

The 24 FS which were selected were relevant to students' writing assignment as it is expected to consist of an introduction, a literature survey, a section focusing on the aim and research questions, and finally a method section. All 24 FS consisted of at least two content words and occurred in the COCA (Davies, 2008-), in which they had a frequency of 30 or above. Next, the target items were divided into three groups (see also Treatment): 1) target items which were only offered in awareness-raising or recognition activities, 2) target items which were only offered in cued output activities, and 3) target items which were offered in both recognition and cued output activities. Frequency of occurrence was held constant for all items, i.e. all FS occurred four times in the learning materials that were provided. Regarding

the items that were assigned to the Recognition + Cued Output activities, they were offered twice in recognition activities and twice in cued output activities. However, the items were not counterbalanced for type of activity (recognition, cued output, or both) because this was not feasible in the context of an authentic classroom setting.

All target items were piloted in the academic year before. As the target items had to be relevant to the students' end-of-year-assignment, ecological validity was given high priority in the selection procedure. As a result, some target items were synonymous (*findings/results/studies reveal/indicate/demonstrate*), while others contained the same verb or noun (*conduct an interview, conduct an experiment* or *recent evidence suggests, the evidence is inconclusive*).

Treatment

The treatment was organized in three consecutive weeks in the second semester of the academic year. In other words, three learning sessions were mainly devoted to academic FS. The learning materials that were used in the treatment consisted of two types of activities which were specifically developed for this study: recognition activities and cued output activities. The activities were designed according to the principle of form-focused vocabulary instruction (Laufer, 2005) or explicit vocabulary learning (Schmitt, 2008) since it has been argued that explicit instruction is an optimal approach for (initial) form-meaning mapping in classroom-based courses (Hulstijn, 2003; Schmitt, 2008). In addition, several studies have shown that explicit instruction is beneficial for learning FS (Laufer & Girsai, 2008; Peters, 2014; Webb & Kagimoto, 2011). All activities were piloted in the year before.

The recognition activities aimed at raising learners' awareness of the (omni-)presence of formulaic sequences in academic writing and are in line with Lewis' (1993) lexical approach. Three types of activities were designed. In a first type of activity, learners were provided with

sentences or excerpts from scholarly articles in which they had to underline FS that could be considered relevant to academic writing. The excerpts obviously contained more FS than the ones that were targeted. These were then also briefly discussed in class.

Example 1:

- Although there are numerous other empirical studies which have investigated the linkage between the stock market and the options market, *the evidence is* still *inconclusive*.
- *A considerable amount of* literature has shown that extroverts make the best leaders. These people are favored in hiring and promotion decisions, and they are perceived to be more effective by supervisors and subordinates alike. To test this idea, we conducted a field study, in which, we sent questionnaires to managers and employees at 130 franchises of a U.S. pizza delivery company. We asked bosses to rate how extroverted they considered themselves, and asked employees to estimate how often they and their colleagues "try to bring about improved procedures" among other proactive behaviors. *We collected data* on each store's profitability, controlling for variables such as whether the franchise was in a high-volume college town.

In a second type of activity, learners had to recognize the more academic sentence of two sentences.

Example 2:

- They found the same things as Waters (2002).
- Their findings were consistent with Waters (2002).

In a third type of activity, learners were asked to indicate in which part of their paper (1) Literature, 2) Aim, or 3) Method) they would use the following sentences.

Example 3:

- A central issue in marketing research has been the effect of above-the-line promotion.
- Our questionnaire consists of nine Likert-scale questions.
- A qualitative approach was adopted.

The cued output activities consisted of *Fill in the gap* exercises (sentences or text excerpts, see example 4), *Rephrase* activities with or without a clue (in isolated sentences or in short excerpts from scholarly articles, see example 5), and *Use in a sentence* activities (see example 6). In all these activities, learners were always asked to provide the intact FS and not to combine the individual constituents making up the FS in order to prevent them from making infelicitous combinations (see also Peters, forthcoming; Durrant & Mathews-Aydınlı, 2011) since formulaic sequences are

> best learned as integral wholes or independent entities, rather than by the process of placing together their component parts, either because (a) they may not be understood or appropriately produced without specific knowledge, or (b) because they occur with sufficient frequency that their independent learning will facilitate fluency (Durrant & Mathews-Aydınlı, 2011, p. 60).

Example 4: Fill in the gap (students could choose FS from a list)

- Qualitative in-depth interviews were used to study consumer roles in different service settings. Each …… (= *interview was done*)

by the principal researcher and lasted between thirty and forty-five minutes. (*interview was conducted*)

- The study …….(= *wants to find out/research*) how women leaders experience their roles as leaders. (*seeks to investigate*)

Example 5: Rephrase

- **We do not know a lot about** the effect of below-the-line promotion. (*There has been little research into ...*)
- In recent years, there has been **a lot of research** investigating Word-of-Mouth in online space. (*a considerable amount of research*)

Example 6: Use in a sentence

- to administer a survey:
- to obtain data:

The range of activities was intended to reflect the range of activities proposed by Nation (2013, pp. 103-112) to foster lexical learning: noticing, retrieval, and creative use. The recognition activities aimed at "noticing" the FS, the gap filling and rephrasing activities correspond to "retrieval", while the use in a sentence activity can be considered a limited form of creative use. Each learning session comprised recognition as well as cued output activities. As one encounter is unlikely to be sufficient to create a strong form-meaning link in the mental lexicon (Hulstijn, 1992) and repetition in vocabulary-focused activities is beneficial for word learning (Laufer & Rozovski-Roitblat, 2011; Nation, 2013; Peters, 2014), the target items were offered four times in the exercises. Note that eight target items were only offered

in recognition activities, eight items only in cued output activities, and eight items in both types of activities (see also Target Items).

Tests

Given the incremental nature of vocabulary learning, it was considered important to measure learning gains at different levels of sensitivity. Therefore, three types of tests at different levels of sensitivity were developed in order to get a more accurate picture of students' learning gains: a recognition test, a cued output test, and a writing test. Nation and Webb (2011) argue that recall tests are more difficult than recognition tests because in the former the form of the lexical item needs to be retrieved from memory, which "requires substantial strength of knowledge in order to answer questions correctly" (p.304). The three tests were used as pretest and as posttest but the order of the items in the cued output test and of the excerpts in the recognition test was changed to control for any order effects. All tests were piloted in the academic year before. In addition to the writing tests, students' spontaneous use of FS in their end-of-year assignment was also analyzed.

In the recognition test, students were provided with several excerpts from scholarly articles in which they had to underline academic FS. The excerpts could contain other FS but these were not taken into account in the analysis, which was confined to the target items.

Example 1

- Based on the results of both stages of measurement, several conclusions can be drawn. Each of these is discussed below. First, findings indicate that spectators largely attend the tournament in groups (98.9%), either with friends, family, or both. These results are consistent with previous research documenting the importance of groups to sport event spectators.

In the cued output test, students were required to supply the FS as a whole. The Dutch translation was provided in brackets. The first letter of each word of the FS was also provided in order to prevent students from supplying alternative but correct answers.

Example 2

- **D……** for the study was **c……..…** (= *data werden verzameld*) during the 2003-2004 and 2004 –2005 academic years from students enrolled in Intermediate Accounting at a state university located on the East Coast. (*Data ... collected*)

The writing test consisted of a reading and writing part. First, students read an interview (spoken register) in which researchers explain a study they have recently conducted (rationale, research questions, method, results). Next, learners summarized the interview in a written, academic register. The summary was expected to contain 250-350 words. A framework with guidelines on which information to include was provided. The writing test was piloted to verify whether it was indeed possible to use a considerable number of the target items. Although it was obviously not possible to include all 24 target items in the summary, especially since some FS were synonymous (*findings reveal* or *findings indicate*), the pilot showed that 22 FS could potentially be used in the summary.

Finally, students' end-of-year assignments were also analyzed in terms of the number of FS used. As has already been mentioned, students had to write a collaborative research paper consisting of an introduction, a literature section, an aim and research questions section, and a method section. Although students could do this assignment at home in pairs or groups of three and use websites and dictionaries, it was nevertheless considered worthwhile to

investigate the spontaneous use of FS in their writing. Moreover, their papers were compared with papers written by students who had not attended the sessions on academic FS[4].

Procedure

The study was carried out at the beginning of the second semester in two parallel groups of second-year business English students during regular class time. An English class normally takes two hours. The study consisted of five sessions. In the first week, students took the pretests: first the writing test, next the cued output test, and finally the recognition test. Students were told that these tests were administered to verify their familiarity with academic English. However, they were not informed that they would have to take the same tests in week 5. In the second, third and fourth week of the second semester, students engaged in vocabulary-focused activities on FS. In each session, they performed recognition as well as cued output activities. Students were told that these exercises aimed at improving their academic English for the collaborative research paper they were required to write. They were also familiarized with the *Phrasebank of the University of Manchester* where they could find more examples of FS. The activities in the three sessions took approximately 2 hours and 45 minutes, which means that a session was never completely devoted to the academic FS. In the fifth week, students were administered the same tests as in week 1 without any warning. However, the order of the items in the cued output test and the order of the excerpts in the recognition test were changed. As in week 1, students first took the writing test, then the cued output test, and finally the recognition test. They were also debriefed about the aim of the study. In week 11 of the second semester, students had to submit their collaborative research paper. Table 2 summarizes the procedure.

Table 2

Procedure

When	What
Session 1 (week 1)	Pretests • Writing test • Cued output test • Recognition test
Session 2 (week 2)	Recognition activities Cued output activities
Session 3 (week 3)	Recognition activities Cued output activities
Session 4 (week 4)	Recognition activities Cued output activities
Session 5 (week 5)	Posttests • Writing test • Cued output test • Recognition test
Week 11	Submitting collaborative research paper

Scoring and analyses

The pretests and posttests were scored as follows:

Recognition test:

- 0 when students did not underline the FS or underlined only one word (and not a sequence of words), e.g. *findings* or *approach*. Articles and auxiliaries were not taken into account, e.g. *the findings* was also scored as incorrect.
- 0.5 when students did not underline the complete FS, e.g. not the preposition (*the following findings emerged*).
- A score of 1 was awarded when student underlined the complete FS, e.g. *the following findings emerged from*.

Cued output tests:

- Students did not receive a score 1) when they did not supply an answer, 2) when one or more of the content words were incorrect or missing (*a certain important* instead of *a central issue*; *the evidence* instead of *the evidence was inconclusive*).
- A score of 0.5 was given 1) when the FS contained a spelling mistake (*objectif, seaks*), 2) when the preposition was incorrect or missing (*little research* or *little research to*), or 3) when one of the words constituting the FS was incorrect but formally similar to the target word (*the main object* instead of *the main objective*).
- A score of 1 was awarded when the FS as a whole was correct. Morphological and grammatical mistakes such as subject-verb agreement were not taken into account. For instance, *the results indicates* was considered correct.

Writing tests/collaborative research papers:

Students received a score of 0.5 for FS that were misspelled (*adressed*) or that contained an incorrect preposition. A score of 1 was awarded for FS that were completely correct. Incorrect attempts were also counted but students did not receive any score for those items, e.g. *they addressed a survey to* instead of *to administer a survey* was considered an incorrect

attempt. In addition to the writing tests, we also analyzed students' use of the 24 selected target FS in their collaborative research papers. For this, the number of FS used (tokens) in papers written by students who had attended two or three sessions on academic FS was compared with the number of FS used in papers written by students who had not attended these sessions.

All tests were scored by two raters. The interrater reliability for the recognition pretest, as indicated by Pearson Product Correlation, was $r = .99$ ($p < .0001$), for the recognition posttests $r = .99$ ($p < .0001$), for the cued output pretest $r = .99$ ($p < .0001$), for the cued output posttest $r = .98$ ($p < .0001$), for the writing pretest Spearman's rho = .75 ($p < .0001$)[5], and for the writing posttest Pearson $r = .93$ ($p < .0001$). The items on which the two raters differed were discussed and a consensus was reached.

To determine whether the participants had made any learning gains from pretest to posttest, paired t-tests were employed for each test type. In case, the assumptions of normality were not met, non-parametric analyses were computed (*Wilcoxon matched-pair signed-rank* test). A *Mann-Whitney U* test was computed to analyze students' collaborative research papers. To determine whether type of instruction had an effect on students' learning gains, Friedman tests were employed.

Results

Results related to the first research question

The first research question addressed the effect of vocabulary-focused instruction on students' learning gains, measured at three different levels of sensitivity: recognition, cued output, and spontaneous use. First, the results for each test type will be presented. The results section ends with an analysis of learners' end-of-year assignments.

Recognition tests. Participants were able to recognize about one fourth of the FS in the recognition pretest (see Table 3). Their recognition of FS increased by 52%, which is approximately 9 FS. The increase was statistically significant, as indicated by a paired *t*-test (t = -8.40; df = 26; p < .0001). Two students were not able to finish the recognition posttest. As a result, their scores were not taken into account in this analysis.

Table 3

Mean, Minimum and Maximum Scores of Pretests and Posttests (Standard Deviation in brackets)

	Pretest (Max = 24)			**Posttest (Max = 24)**			**Absolute gains (Max = 24)**			**Relative gains (100%)**		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max			
Recognition	6.79 (4.25)	0	14.5	15.65 (5.87)	0	24	8.87 (5.49)	0	18	51.67 (30.61)	0	100
CuedOutput	8.36 (3.15)	2	15	14.26 (3.51)	5	21.5	5.90 (3.15)	1	18	26.80 (14.34)	4.55	73
WritingToken	0.10 (0.31)	0	1	2.88 (2.00)	0	7.5	2.78 (1.98)	0	7.5	11.62 (8.34)	0	31
WritingType	0.10 (0.31)	0	1	2.72 (2.00)	0	7.5	2.62 (1.90)	0	7.5	13.22 (10.67)	0	46

Cued output tests. Participants were able to provide on average almost six FS more on the cued output posttest compared to the pretest, which is an increase of almost 27%. The difference between the pretest and the posttest was statistically significant (t = -10.07; df = 28; p < .0001).

Writing tests. As can be seen in Table 3, only a very limited number of students (N = 3) used a FS in the writing pretest. The average increase from pretest to posttest was 2.78

(11%) when all FS (tokens) were counted and 2.62 (13%) when only FS types were counted. The *Related-Samples Wilcoxon Signed Rank* Tests indicated that the increase in the number of FS used was statistically significant in the case of FS tokens ($p < .0001$) as well as in the case of FS types ($p < .0001$). Twenty-one of the 24 FS[6] were used in students' writing posttests, which shows that it was indeed possible to use the FS in the writing test.

Collaborative research papers. In addition to the writing tests, we also analyzed students' end-of-year assignments to shed more light on students' spontaneous use of FS (tokens). In total, 24 papers (vocabulary-focused instruction) were compared with 17 papers (no vocabulary-focused instruction). Although from a pedagogical perspective students' spontaneous use of FS in their papers is the most relevant, the results of this analysis should be interpreted with care since students could use their learning materials, the website of the *Phrasebank from the University of Manchester*, or dictionaries.

As can be gleaned from Table 4, students who had received explicit instruction on academic FS used on average 10 FS in their paper, whereas students who had not received such instruction used only 1.5 FS. The difference was statistically significant, as indicated by a *Mann-Whitney U*-test ($p < .0001$). All 24 FS were used in the papers written by students who had received FS instruction, whereas the FS used in papers written by students who had not received explicit instruction on FS were confined to *a considerable amount of*, *questionnaire consists of*, *conduct an interview*, *the main objective of*, *the paper focuses on*, and finally *obtain data*.

In sum, most learning gains were to be found at the level of recognition. Students were also able to provide more FS in the cued output posttest than in the cued output pretest but the learning gains were lower compared to the recognition tests. Finally, they spontaneously used more FS in their writing in the writing posttest as well as in their collaborative paper.

Table 4

Mean, Minimum and Maximum number of FS used in collaborative paper (standard deviation in brackets)

	N	Mean	Min	Max
+ FS instruction	24	9.98 (6.27)	0	28
- FS instruction	17	1.53 (2.38)	0	8

Results related to the second research question

This study also sought to shed some light on the effect of type of instruction: recognition activities aiming at raising students' awareness/recognition of FS, cued output activities aimed at increasing form recall of FS, or a combination of both (see Table 4). Non-parametric analyses (*Friedman* tests) were conducted since the assumption of normality was not met for each subgroup of items. The analysis was based on students' relative gains to take into account the FS students were already familiar with. The relative gains were computed as follows:

Relative gains:

(Score Posttest – Score Pretest / 8 – Score Pretest) x 100 =

(5 – 3 / 8 – 3) x 100 = 40%

Recognition test. Regarding the research question whether type of instruction has an effect on students' recognizing FS, the answer is positive (see Table 5; *Chi*-Square = 9.74; df = 2; p

= .008). The increase in learning gains was significantly higher for the Cued output items than for the Recognition items (p = .02).

Cued output test. In addition, type of instruction also had a significant effect on the cued output test (*Chi*-Square = 7.49; df = 2; p = .02). Only the Recognition+CuedOuput items differed significantly from the Recognition items, as indicated by a *Related Samples Wilcoxon Signed Rank* test (p = .001) (see also Table 5).

Table 5

Mean scores per subgroup of target items and per test type (standard deviation in brackets)

	Recognition test (N = 27)		**Cued output test** (N = 29)		**Writing test** (N = 29)	
	Pretest (Max = 8)	**Posttest** (Max = 8)	**Pretest** (Max = 8)	**Posttest** (Max = 8)	**Pretest** (Max = 8)	**Posttest** (Max = 8)
Recognition items	2.87 (1.68)	4.98 (1.96)	2.26 (1.27)	3.85 (1.46)	.07 (.27)	.59 (.79)
Cued output items	1.94 (1.57)	5.56 (2.07)	3.17 (1.345)	5.20 (1.28)	.07 (.39)	1.61 (1.34)
Recognition+Cued output items	1.94 (1.70)	5.06 (2.33)	2.96 (1.47)	5.39 (1.68)	.00 (.00)	.56 (.80)

Writing test. Finally, type of instruction also affected the number of FS students used in the writing posttest (*Chi*-square = 12.91; df = 2; p = .002). A *Related Samples Wilcoxon Signed Rank* test indicated that students used significantly more Cued output items than Recognition items (p = .022) in the writing posttest. No other differences were found.

Given the limited number of FS per type of item (k = 8) and the fact that the three groups of target items were not counterbalanced, the results presented here should be

interpreted with caution. Caution should also be exercised when interpreting the findings of the writing posttests because students may have considered some FS more relevant than others. The effect of type of instruction on the FS used in students' end-of-year assignment was not analyzed because this analysis was not deemed relevant since students could use all learning materials and the website of the *Academic Phrasebank* while writing their paper.

Table 6

Absolute and relative gains per test type and per type of instruction (Standard deviation in brackets)

	Recognition Items		**Cued Output Items**		**Recognition+Cued output Items**	
	AbsGains Max = 8	RelGains %	AbsGains Max = 8	RelGains %	AbsGains Max = 8	RelGains %
Recognition test	2.11	41.65	3.61	58.27*	3.11	49.21
	(1.86)	(36.91)	(2.27)	(37.45)	(2.38)	(39.16)
Cued Output test	1.64	27.80	1.86	35.56	2.43	48.23*
	(1.25)	(21.30)	(1.63)	(30.25)	(1.65)	(28.01)
Writing test	0.66	8.19	1.50	18.75*	0.64	7.97
	(0.80)	(10.04)	(1.36)	(16.95)	(0.83)	(10.42)

Note: AbsGains = absolute gains; RelGains = relative gains

Discussion

This study aimed to explore whether a vocabulary-focused approach to academic FS would result in significant learning gains. The study also attempted to shed some light on the effect of type of instruction on any observed learning gains. The findings suggest that a vocabulary-focused approach to academic FS indeed results in significant learning gains at the level of

recognition, cued output and spontaneous use of academic FS. Moreover, students were able to transfer their increased level of awareness and accurate (cued) output to their end-of-year assignment. The findings regarding the type of instruction are not conclusive but it seems that activities involving cued output exercises might be more beneficial not only at the level of production but also at the level of recognition.

The vocabulary-focused treatment adopted in this study comprised two of the three psychological conditions advocated by Nation (2013) that may result in lexical learning: noticing and retrieval. The third process "creative use" was only offered minimally in the treatment. Nation stresses the importance of students' engagement with new words. Although Nation advocates these three conditions for single words, previous research has shown that they might be relevant to learning FS as well (Peters, 2012, 2014, forthcoming). It can be argued that students' involvement with the learning materials and the awareness-raising activities in particular was high since these tasks were relevant to their learning goals, viz. writing a collaborative research paper. The activities also made the target items salient. Not only did students have to recognize academic FS in text excerpts, they also engaged in activities in which the FS were decontextualized. For instance, differences with more colloquial English were pointed out. Students were also taught in which section of their paper they could use the FS. The cued output activities provided the students with retrieval opportunities, which created the opportunity to strengthen the memory trace to the form of the target item. Although our vocabulary-focused treatment only comprised recognition and cued output activities and did not even take three hours, students not only exhibited learning gains when they were prompted to recognize or provide FS but they were also able to use FS spontaneously in the writing posttests. Moreover, a considerable number of students also used the FS in their end-of-year assignment. Our findings suggest that explicit, vocabulary-focused

instruction has the potential to boost students' knowledge of FS and are thus in line with findings from previous studies (Jones & Haywood, 2004; Li & Schmitt, 2009).

Overall, the findings of this study support the arguments of previous researchers advocating explicit instruction in FS (Coxhead & Byrd, 2007; Meunier, 2012). The importance and necessity of such instruction was particularly illustrated in the results of the pretests. Although students were able to recognize on average 7 FS and to provide FS when prompted, they barely used any FS in their writing pretest. In addition, students who had not received vocabulary-focused instruction on FS used significantly fewer FS in their end-of-year assignment. In general, students' writing posttests and end-of-year assignments conformed more to academic expectations and were less colloquial in style. The use of academic FS helped students to meet the expectations of the academic community, to signal stages in their discourse, and to write in an appropriate, non-colloquial, formal register. Let us illustrate this with the example *to administer a survey*. Students who had not received FS instruction drew on their L1 and produced sentences in their paper, which contained literal translations from Dutch *afnemen* (*to take*), such as:

- A Survey will be taken with the students who are following the architecture education.
- We would take a survey on graduating students.

However, students who had engaged in vocabulary-focused activities tended to write sentences in their paper such as:

- To facilitate this process, 'e5mode' can administer a survey to their customers.
- This may be done by administering a survey to both chefs of a business unit and employees.
- This questionnaire will be administered to European consumers from all ages, professions and countries.

Given the limited number of target items per instruction type and the fact that the three instruction types were not counterbalanced, the results related to the effect of type of instruction should be interpreted with great care. However, the findings tentatively suggest that activities involving cued output (or retrieval) tasks might be more beneficial than tasks aiming only at awareness and recognition. These results are consistent with those of Webb (2005) who found that productive vocabulary learning tasks might be more beneficial for single word learning than receptive tasks when time-on-task is not controlled for. Our findings also seem to corroborate Boers and Lindstromberg's (2012) argument that awareness-raising may not be such "a powerful accelerator" (p. 99) of learning after all. It may also explain why Jones and Haywood (2004) did not find clear evidence for the beneficial effects of vocabulary-focused instruction on cued output and spontaneous use of FS since their treatment mainly comprised awareness-raising activities. Although their writing component did contain some fill-in-the-gap exercises, students were merely encouraged to use FS and not to actually retrieve FS as was the case in the present study. It should be pointed out here that there are, obviously, other ways of fostering FL learners' awareness and use of FS (e.g. using concordances) and of selecting appropriate FS (e.g. the *Academic Formulas List*, developed by Simpson-Vlach & Ellis, 2010; the *Academic Collocation List* (Ackermann & Chen, 2013)) than the ones used in the present study.

There are a number of limitations in this study. As has already been mentioned, the target items in the three types of instruction (recognition, cued output, combination of recognition and cued output) were not counterbalanced. Although almost all FS were used by at least one student in the writing posttest, some FS may nevertheless have been more relevant to the writing assignment than others. A third limitation concerns the lack of a control group. Although a pretest-posttest within-subject design was adopted and significant learning gains were found, the study would have benefited from a control group that would only have taken

the tests. However, since the data were collected in a real classroom setting, it was not feasible to operationalize such a design.

Conclusion

The findings of this study suggest that explicit, vocabulary-focused instruction on academic FS has the potential to boost students' awareness, cued output and spontaneous use of FS and should thus be included in any course on academic English or academic writing. The combination of (decontextualized) awareness-raising and recognition activities, cued output activities and repetition proved to be fruitful in having students engage repeatedly and thoroughly with the target items. Given the limited amount of time that was devoted to familiarizing students with these academic FS, the results could be considered particularly positive.

Endnotes

1. Cortes uses the term lexical bundles, which she defines as "combinations of three or more words that frequently occur in a language or a given register (p. 34).

2. There are four parallel groups in total but only two groups were involved in this study. Author 1 was also the lecturer of these two groups.

3. This means that our target items were not lexical bundles and were not solely based on frequency data.

4. Only two of the four parallel groups participated in the study but their end-of-year-assignments were compared with those written by students in the other two groups who had not performed the activities on academic FS.

5. A Spearman rho correlation was used because the data were not normally distributed.

6. The following FS were not used by any student: *the questionnaire consists of …*, *to select participants*, and *to adopt an approach*.

References

Ackermann, K. & Chen, Y. (2013). Developing the Academic Collocation List (ACL) – A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes, 12*, 235-247.

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English. *English for Specific Purposes, 31*, 81-92.

Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, *14*(3), 275–311. doi:10.1075/ijcl.14.3.08bib.

Boers, F., J. Eyckmans, J. Kappel, H. Stengers, & M. Demecheleer (2006). Formulaic sequences and perceived oral proficiency: putting a lexical approach to the test. *Language Teaching Research, 10* (3), 245-261.

Boers, F., & Lindstromberg, S. (2005). Finding ways to make phrase learning feasible: the mnemonic effect of alliteration. *System, 33*, 225-238.

Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics, 32*, 83-110.

Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*, 30-49.

Cortes, V. (2013). The purpose of this study is to: connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes, 12*, 33-43.

Cowie, A. (1992). Multiword Lexical Units and Communicative Language Teaching. In P. Arnaud & Béjoint, H. (eds.), *Vocabulary and Applied Linguistics* (pp. 1-12). London, MacMillan.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213-238

Coxhead, A. (2008). Phraseology and English for academic purposes. Challenges and opportunities. In F. Meunier & S. Granger (eds.), *Phraseology in language learning and teaching* (pp. 149-161). Amsterdam. John Benjamins.

Coxhead, A. & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing, 16*, 129-147.

Davies, M . (2008 –). *The Corpus of Contemporary American English (COCA): 450 million words, 1990–present*. Available online at http://corpus.byu.edu/coca/ .

Durrant, P., & Mathews-Aydınlı, J. (2011). A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes, 30*, 58-72.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text, 20,* 29–62.

Gilquin & Paquot, M. (2008). Too chatty. Learner academic writing and register variation. *English Text Construction, 3* (1), 41-61

Hulstijn, J.H. (1992). Retention of inferred and given word meanings: experiments in incidental vocabulary learning. In P. L. J. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113–125). London: Macmillan Academic and Profession Ltd.

Hulstijn, J. H. (2003). Incidental and Intentional Learning. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 349–381). Malden, MA: Blackwell.

Hüttner, J. (2007). *Academic writing in a foreign language. An extended genre analysis of student texts*. Frankfurt am Main: Peter Lang.

Hyland, K. (2008). As can be seen: lexical bundles and disciplinary variation. *English for Specific Purposes, 27*, 4-21.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics, 32*, 150-169.

Jones, M. & S. Haywood (2004). Facilitating the acquisition of formulaic sequences: an exploratory study in an EAP context. In N. Schmitt (ed.), *Formulaic Sequences* (pp. 269-292). Amsterdam: John Benjamins.

Laufer, B. (2005). Focus on form in second language vocabulary learning. In S. H. Foster-Cohen, M. Garcia-Mayo, & J. Cenoz (Eds.), *Eurosla Yearbook. Vol. 5* (pp. 223-250). Amsterdam: John Benjamins Publishing Company.

Laufer, B. & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: a case for contrastive analysis and translation. *Applied Linguistics, 29*, 1-23.

Laufer, B. & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research, 15*, 391-411.

Lewis, M. (1993). *The Lexical Approach. The State of ELT and a Way Forward*. Language Teaching Publications.

Li, J. ,& Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: a longitudinal case study. *Journal of Second Language Writing, 18*, 85-102.

Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: a multi-corpus study. *English for Specific Purposes, 31*, 25-35.

Lindstromberg, S. & Eyckmans, J. (forthcoming). How big is the positive effect of assonance on the recall of L2 collocations. *ITL, International Journal of Applied Linguistics*.

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics, 32*, 111-129.

Nation, I.S.P. (2013). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston: Heinle.

Peters, E. (2009). Learning collocations through attention-drawing techniques : a qualitative and quantitative analysis. In: Barfield A., Gyllstad H. (Eds.)*, Researching collocations in another language: multiple interpretations*. Basingstoke: Palgrave MacMillan, 194-207.

Peters, E. (2012). Learning German formulaic sequences: the effect of two attention-drawing techniques. *Language Learning Journal, 40* (1), 65-79.

Peters, E. (2014). The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations. *Language Teaching Research, 18* (1), 75-94.

Peters, E. (forthcoming). The learning burden of collocations: the role of interlexical and intralexical factors. *Language Teaching Research*.

Schmitt, N. (2008) Review article: instructed second language vocabulary learning. *Language Teaching Research, 12*, 329–63.

Schmitt, N. & R. Carter (2004). Formulaic sequences in action: an introduction. In N. Schmitt (Ed.), *Formulaic Sequences* (pp.1-22). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Simpson-Vlach, R., & Ellis, N. (2010). An academic formulas list: new methods in phraseology research. *Applied Linguistics, 31*, 487-512.

Sinclair, J. (1991). *Corpus, concordance, collocation: Describing English language*. Oxford: Oxford University Press.

Webb, S. (2005). The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27*, 33-52.

Webb, S., & Kagimoto, E. (2010). Learning collocations: do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics, 32*, 259-276.

Webb, S., Newton, J., & Chang, A.C.S. (2013). Incidental learning of collocation. *Language*

*Learning, 6*3, 91-120.